

WE CLAIM:

1. A system for discovering and monitoring an active topology of a dual counter rotating ring (DCRR) at a data link layer in order to enable optimal frame forwarding from a medium to the DCRR and to provide rapid failover in a network component failure event, each access point in the DCRR comprising:

a) a topology maintenance protocol (TMP) entity for monitoring individual links of the DCRR, discovering the active topology of the DCRR, communicating topology changes to other access points in the DCRR, and informing frame forwarding and learning processes of the topology to enable failover to a redundant path;

b) a frame forwarding process for selecting a shortest path route between a source and destination access point in the DCRR; and

c) an address learning process for inspecting a source medium access control (MAC) address in each received frame to learn a shortest path route for each MAC address.

2. A system as claimed in claim 1 wherein the topology maintenance protocol entity comprises:

a) a Receive PDU Processor for receiving incoming protocol data units (PDUs), extracting information from the PDUs and distributing the information to other local topology learning entity components, as required;

b) a link integrity monitor machine (LIMM) for deciding a sanity of links terminated at the access point, and communicating the state of links terminated at

the access point to other local topology learning components and to upstream link partners;

c) a reconfiguration machine (RM) to determine DCRR topology parameters required by the forwarding process;

d) a topology discovery machine (TDM) for determining an active topology of the DCRR;

e) a reconfiguration notification machine (RNM) for controlling insertion of ring reconfiguration messages into a ring of the DCRR;

f) a link integrity PDU transmit (LIPT) machine for controlling insertion of link integrity PDUs onto a ring of the DCRR; and

g) a PDU insertion machine (PIM) which arbitrates with the tandem datapath payload queue for insertion of the PDUs onto the DCRR.

3. A system as claimed in claim 1 wherein the receive PDU processor extracts protocol information from received PDUs, passes the information to other TMP entity components and terminates one or more predetermined PDU message types.

4. A system as claimed in claim 3 wherein the PDU message types terminated by the receive PDU processor comprise:

a) link path integrity messages; and

b) if the source address of a message matches an address of the access point:

i. link failure messages;

ii. link recovery messages; and

iii. topology discovery messages.

5. A system as claimed in claim 4 wherein each topology maintenance protocol entity includes two receive PDU processors, one for each WAN port of a respective access point.

6. A system as claimed in claim 1 wherein the PDUs comprise:

a) link path integrity messages (LPIM) used to determine the state of a bi-directional link between two adjacent access points in the DCRR;

b) link failure messages (LFM) used to communicate the failure of network links to other access points in the DCRR;

c) link recovery messages (LRM) used to communicated the recovery of network links to other access points in the DCRR;

d) link topology discovery messages (TD) used to discover the topology of the DCRR.

7. A system as claimed in claim 6 wherein information in the LFM comprises:

a) a frame type identifier;

b) a source link status containing a current state of the source link; and

c) an integrity checksum.

8. A system as claimed in claim 6 wherein information in each of the LFM, LRM and TDM comprises:

a) a frame type identifier;

b) a time to live variable;

c) a source address of an MAC that initiated the LRM onto the ring; and

d) an integrity checksum.

9. A system as claimed in claim 2 wherein the L IMM monitors link path integrity messages sent by an upstream link partner access point to determine a state of the link path and provides the downstream state of the link path to the upstream link partner access point.

10. A system as claimed in claim 9 wherein the topology protocol entity includes two instances of the L IMM, one for each ingress link in the DCRR.

11. A system as claimed in claim 9 wherein the L IMM further performs the functions of:

a) periodically passing to an opposite port's RNM a link failure assertion message when the L IMM is in an operational state;

b) periodically passing to the RM a link operational level assertion message when the L IMM is in an operational state;

c) periodically passing to the PIM a local link state message which represents a current state of the L IMM.

12. A system as claimed in claim 9 wherein the L IMM receives from a management function the following operating parameters:

a) a nominal time between successive transmissions of LPIMs;

b) a number of LPIMs required to cause a transition from a non-operational state to an operational state;

c) a number of LPIM receipt periods that must elapse to cause a transition from an operational state to a link fail state; and

d) a number of LPIMs required to cause a transition from the link fail state to the operational state.

13. A system as claimed in claim 2 wherein the RM provides the forwarding process with information about the status of the DCRR and records the reception and generation of LFM and LRM messages at each DCRR port to control forwarding and learning behavior between reception of the messages.

14. A system as claimed in claim 13 wherein the RM further performs the actions of:

- a) initiating a topology discovery process;
- b) enabling port ADD and TANDEM payload datapaths;
- c) informing the forwarding and learning processes of the topology of the DCRR; and
- d) transitions to a disconnected state if both access point ports are non-operational as indicated by the LIMMs.

15. A system as claimed in claim 2 wherein the TDM performs topology discovery under the control of the RM and maintains a timer for the sending of topology discovery messages.

16. A system as claimed in claim 15 wherein the TDM further performs the following functions:

a) informs the RM of a completion of a topology discovery procedure, and the configuration of the DCRR on completion of the topology discovery procedure;

b) sends topology discovery message requests to the TPIM when a topology discovery message is to be sent or an acknowledgement of a topology discovery message received from another access point in the DCRR is to be sent;

c) informs management of the occurrence of a timeout while waiting for a second topology discovery message during a topology discover procedure.

17. A system as claimed in claim 2 wherein the RNM registers when a link failure or a link recovery event occurs at an opposite DCRR port, and maintains a hysteresis timer on a state of an opposite DCRR port link to permit the port link to settle in the event of a link failure or link recovery.

18. A system as claimed in claim 17 wherein there is one instance of the RNM for each DCRR port.

19. A system as claimed in claim 17 wherein the RNM receives from a management function the following parameters:

a) a value representing a minimum time period for permitting a link to settle between the transmission of an LRM and the transmission of an LFM; and

b) a value representing a minimum time period for permitting a link to settle between the transmission of an LFM and the transmission of an LRM.

20. A system as claimed in claim 2 wherein the LIPT is a free running timer scaled to achieve the fastest possible failure discovery while not conflicting with Layer 1 transport protection mechanisms.

21. A system as claimed in claim 20 wherein the LIPT receives from a management function a value representative of a time period which controls a frequency of transmission of link path integrity messages.

22. A system as claimed in claim 2 wherein therein the is an instance of the PIM for each DCRR port.

23. A system as claimed in claim 22 wherein each PIM arbitrates for access to a corresponding link on a frame-by-frame basis with PDU insertion as a highest priority, TANDEM path access has second priority, and ADD path access has lowest priority.

24. A system as claimed in claim 1 wherein the forwarding process performs a hash over the address of multicast and flooded traffic to yield a single bit has result that is used to specify a direction for forwarding around the DCRR.

25. A system as claimed in claim 24 wherein IP addresses are further included in the hash calculation to ensure load sharing between routers and between routers and servers subtending the traffic.

26. A system as claimed in claim 1 wherein when an address of a frame is known, the forwarding process

chooses a shortest path direction for forwarding a frame by learning a number of hops to another given access point that subtends the traffic.

27. A system as claimed in claim 1 wherein if the DCRR is a linear or broken ring topology, the forwarding process forwards all flood and multicast frames in each direction.

28. A method for discovering and monitoring an active topology of a dual counter rotating ring (DCRR) at a data link layer of the DCRR in order to enable optimal frame forwarding from a local area network (LAN) to the DCRR and to provide rapid failover in a network component failure event, comprising the steps of:

a) monitoring individual links of the DCRR by sending and receiving link path integrity messages, the link path integrity messages being sent to downstream link partner access points and received from upstream link partner access points, the monitored links being the links on which link path integrity messages are received at each access point in the DCRR;

b) discovering an active topology of the network by periodically entering an active topology discovery state in which topology discovery messages are sent on the ring and topology discovery message responses are received;

c) communicating changes in the active topology due to failure of monitored links to other access points in the network; and

d) informing frame forwarding and address learning processes within each access point of the active

topology to enable failover of communications to redundant resource.

29. A method as claimed in claim 28 wherein the link path integrity messages are sent a rate that ensures rapid failover to redundant resources without interfering with a layer 1 protection protocol of the DCRR.

30. A method as claimed in claim 29 wherein a time interval between link path integrity messages is determined by an administration function that supplies a value representative of a length of the interval.

31. A method as claimed in claim 28 wherein the link path integrity messages are stripped on the downstream end of the monitored link.

32. A method as claimed in claim 29 wherein the link path integrity monitoring is performed by a link path monitoring machine, and there is an instance of the link path monitoring machine at each port of the DCRR.

33. A method as claimed in claim 28 wherein the step of discovering is performed by a topology discovery machine (TDM) that operates under control of a reconfiguration machine (RM) and the TDM discovers the active topology of the DCRR and communicates the active topology to the RM.

34. A method as claimed in claim 33 wherein the RM instructs the TDM to begin a topology discovery process on receipt of a link failure message or a link recovery message, and the TDM reports a topology of the DCRR to

the RM when the link topology discovery procedure is completed, the topology reported being one of: a ring topology and a linear topology.

35. A method as claimed in claim 34 wherein the RM instructs the TDM to reinitiate the topology discovery procedure when the RM receives a link failure message or a link recovery message before an end of the topology discovery procedure.

36. A method as claimed in claims 34 wherein the RM informs the forwarding and learning processes of the topology of the DCRR when the topology discovery procedure is completed.

37. A method as claimed in claim 34 wherein the RM signals to add path queue control logic on transition from one topology to another that a DCRR reconfiguration is in progress and all frames that have been queued at either port must be discarded to avoid frame misordering.

38. A method as claimed in claim 34 wherein the RM signals a payload data enable condition to add path and drop path controllers when the RM is in the ring or the linear topology, and when not enabled the controllers discard all frames that arrive to be added or dropped.

39. A method as claimed in claim 34 wherein the RM signals a tandem path controller when the RM is not in a link breaking or link repairing state and when not enabled the tandem path controller discards all frames arriving for tandeming, the enable signal being sampled frame synchronously.

40. A method as claimed in claim 34 wherein the topology discovery procedure comprises the steps of:

a) requesting a transmit PDU insertion machine at each DCRR port to transmit a topology discovery message;

b) entering a wait state to wait for a return of a first topology discovery message originated by the topology discovery machine;

c) entering a second wait state on receipt of a first topology discovery message to wait for receipt of a second topology discovery message sent by the topology discovery machine; and

d) entering an idle state after receiving the first and second messages and informing the RM of the topology of the DCRR, or

e) entering the idle state after waiting a predetermined time for receipt of the first and second messages and informing the RM of the topology of the DCRR.

41. A method of enabling shortest-path route forwarding of payload data across a communication network comprising two or more access points connected by dual-path links, comprises a pair of link paths adapted to carry payload traffic in a direction opposite that of the other link path, each access point having a pair of WAN ports adapted for connection to respective dual path links, the method comprising the steps of:

a) discovering a topology of the communication network;

b) learning addresses of destination devices subtended from the communication network;

c) determining a path offering a shortest route, in terms of the number of hops to each learned address; and

d) enabling a frame forwarding process to forward payload data across the communication network via the determined path offering the shortest route.

42. A method as claimed in claim 41, wherein the step of discovering a topology of the communication network comprises transmitting a topology discovery message over a selected WAN port of an access point, and waiting a predetermined interval for reception of the transmitted topology discovery message at the other WAN port of the access point.

43. A method as claimed in claim 42, wherein the topology of the communication network is discovered to be a ring topology if the transmitted topology discovery message is received within the predetermined interval, and a linear topology if the transmitted topology discovery message is not received within the predetermined interval.

44. A method as claimed in claim 41, wherein the step of discovering a topology of the network is triggered in response to detection of a change in the topology of the communication network.

45. A method as claimed in claim 44, wherein detection of a change in the topology of the communication network comprises the steps of, in respect of a dual path link connecting first and second access points:

a) transmitting a first link path integrity monitor message from the first access point to the second access point over a first link path of the dual path link;

b) detecting, at the second access point, an integrity of the first link path on the basis of reception of the first link path integrity monitor message;

c) transmitting a second link path integrity monitor message from the second access point to the first access point over a second link path of the dual path link; and

d) detecting, at the first access point, an integrity of the second link path on the basis of reception of the second link path integrity monitor message.

46. A method as claimed in claim 45, wherein the second link path integrity monitor message includes a detected status of the first link path, and the first link path integrity monitor message includes a detected status of the second link path, whereby the first and second access points are informed of the status of a respective downstream link paths.

47. A method as claimed in claim 45, further comprising, when the status of either link path of the dual path link is detected to be in a link failure state, transmitting a link fail message from each of the first and second access points in respective directions away from the dual path link.

48. A method as claimed in claim 47, further comprising, when the status of either link path of the dual path link is detected to change from a link failure state to a link operational state, transmitting a link recovery message from each of the first and second access points in respective directions away from the dual path link.

49. A method as claimed in claim 41, wherein the step of learning addresses comprises the steps of, at each access point:

a) extracting a source address of a payload data frame received at a WAN port of the access point; and

b) maintaining an address cache including data of the source address.

50. A method as claimed in claim 49, further comprising the step of, when the source address of a received payload data frame is not recorded in the address cache, adding data of the source address to the address cache.

51. A method as claimed in claim 50, further comprising the step of flushing the address cache in the event of a detected change in the topology of the communication network.

52. A method as claimed in claim 51, wherein the step of determining a path offering a shortest route comprises the steps of, at each access point:

a) determining a distance, in hops, from a source address of a payload data frame received at a WAN port of the access point; and

b) maintaining an address cache including data of the source address, distance in hops, and WAN port on which the payload data frame was received.

53. A method as claimed in claim 52, further comprising the steps of, in respect of each source address recorded in the address cache:

a) determining whether the distance in hops of a received payload data frame is less than the distance recorded in the address cache; and

b) if the distance in hops of the received payload data frame is less than the distance recorded in the address cache, updating the address cache with data of the distance in hops, and the WAN port on which the payload data frame was received.

54. A method as claimed in claim 53, further comprising the step of flushing the address cache in the event of a detected change in the topology of the communication network.

55. A method as claimed in claim 54, wherein each payload data frame includes a time to live (TTL) value, and wherein determination of the distance in hops to a source address comprises the steps of:

a) setting the TTL to a predetermined value at a source access point;

b) decrementing the TTL at each access point intermediate the source access point and a destination access point; and

c) calculating, at the destination access point, the distance in hops based on a difference between the predetermined value and an actual value of the TTL of the received payload data frame.